Approximating Multiobjective Optimization Problems with Complex Pareto Fronts

Shouyong Jiang
Centre for Computational Intelligence (CCI)
School of Computer Science and Informatics
De Montfort University
The Gateway, Leicester LE1 9BH, U.K.
Email: shouyong.jiang@email.dmu.ac.uk

Shengxiang Yang
Centre for Computational Intelligence (CCI)
School of Computer Science and Informatics
De Montfort University
The Gateway, Leicester LE1 9BH, U.K.
Email: syang@dmu.ac.uk

***Abstract*—The main goal of multiobjective optimization is to achieve a set of well-converged and evenly-distributed Pareto optimal points. While evolutionary algorithms have been reported to converge well, their distribution performance might not be as uniform as we expected, especially when the problems to be optimized involve complex Pareto fronts. In this paper, with the aid of a set of uniformly-distributed reference points, multiobjective optimization problems (MOPs) can be handled by minimizing least reference distances (LRD), which measure the proximity of solutions to their nearest reference points. This way, the uniformity of approximated solutions is implicitly controlled by the reference point set, and convergence is in the charge of LRD. The proposed LRD algorithm (LRDA) is tested and compared with several popular algorithms on a number of old and newly-developed MOPs that have complex Pareto fronts, showing that this method is very promising to obtain evenly-distributed Pareto optimal points for the problems considered in this paper.**

I. Introduction

Multiobjective optimization problems (MOPs) refer to a class of problems having at least two objective functions that conflict one another. Since evolutionary algorithms are able to provide a set of trade-off solutions, known as Pareto-optimal set (POS), without the use of any gradient information of MOPs in question, in a single run, they have been considered as an important method for multiobjective optimization. Over the past several decades, there have been a large number of contributions to multiobjective optimization evolutionary algorithms (MOEAs), e.g. [6], [8], [19], [12], [1], [25], [5], [24], [21], for MOPs.

Generally, convergence and diversity are two fundamental requirements for any well-designed MOEA. The former measures the proximity of approximated Pareto-optimal front (POF) to the true POF whereas the latter emphasizes a good distribution of the approximated POF points. Most often, these two requirements are sort of conflicting, and any improvement in one requirement may lead to a undesirable deterioration in the other. To address this situation, MOEAs are required to have a good balance between convergence and diversity. To date, much effort has been devoted to reaching such a balance by introducing different selection mechanisms, e.g., Pareto-based and decomposition-based selection mechanisms.

Most early-developed MOEAs use Pareto-dominance relations [16] to induce discrimination between two solutions in a population. These MOEAs implement Pareto-based selection in a way that solutions are compared according to their dominance relation and density, where the Pareto-dominance relation is used as the primary selection criterion to promote convergence and the density information, e.g., the crowding distance in nondominated sorting genetic algorithm II (NSGA-II) [5] and the kth nearest neighbour density in strength Pareto evolutionary algorithm 2 (SPEA2) [24], is used as the second selection criterion to maintain diversity. The Pareto-based selection is a first-convergence-then-diversity mechanism, and the second criterion is activated only when the dominance relation cannot make a clear discrimination between solutions. As discussed in [14], such a mechanism has difficulties in meeting special requirements on diversity. In addition, the crowding distance or the kth nearest neighbour density estimation is an approximated diversity measure so that it cannot provide precise density information, which might compromise its ability to produce uniformly-distributed Pareto optimal solutions or even induce some biases (e.g., boundary solutions are favoured over intermediate ones).

In contrast to Pareto-based MOEAs, decomposition-based MOEAs, such as MOEA based on decomposition (MOEA/D) [21] and its variants [14], [13], [15], make use of traditional decomposition methods to convert an MOP into a number of single-objective subproblems and solve them simultaneously in a collaborative manner. Decomposition-based MOEAs work with a set of pre-specified uniformly-distributed weight vectors guiding population members to search along different directions, and the decomposition-based selection procedure involved is carried out by evaluating solutions' achievement on convergence and diversity combined via decomposition approaches. It is believed that an even distribution of weight vector set should result in a set of well-distributed Pareto optimal points. However, some recent studies (e.g., [13], [7], [18]) have revealed that this standpoint is fundamentally flawed, and the use of evenly-distributed weight vectors may not always lead to a uniform distribution of approximated POF points. Even worse, due to the nonlinear mapping property of decomposition approaches, decomposition-based algorithms struggle to distribute Pareto optimal points evenly along complex POFs (e.g., the POF is excessively convex-shaped [13], [18], [11], where a small variation in one objective results in a large gap in another objective).

As mentioned above, Pareto-based MOEAs usually use

Pareto-dominance relations and diversity preservation strategies to meet convergence and diversity requirements whereas decomposition-based MOEAs depend largely on decomposition approaches to obtain a good POF approximation. However, it is often difficult to devise a perfect diversity preservation strategy or decomposition approach to uniformly and precisely distribute Pareto optimal points along the POF. Facing this difficulty, one would naturally think if there is another way to achieve an even distribution of approximated points. Along this direction, in this paper, a novel MOEA based on the least reference distance (LRD), denoted LRDA, is proposed, where LRDA tries to drive a solution toward its nearest reference point or scattered Pareto optimal point, for dealing with MOPs with complex POFs.

The remainder of the paper is organized as follows. Section II presents the motivation of this work and main idea behind the proposed algorithm. The proposed algorithm is detailed in Section III, and experimental studies and comparisons are presented in Section IV. Section V concludes the paper and suggests some future research directions.

II. Motivation

A general mathematical description of an MOP can be given as follows:

$$\begin{aligned} min \quad & F(x) = (f_1(x), ..., f_M(x))^T \\ s.t. \quad & \begin{cases} h_i(x) = 0, & i = 1, ..., n_h \\ g_i(x) \geq 0, & i = 1, ..., n_g \\ x \in \Omega_x \end{cases} \end{aligned} \tag{1}$$

where $\Omega_x \subseteq R^n$ is the decision space, n_h and n_g are the number of equalities and inequalities, respectively, and $F(x)$: $\Omega_x \rightarrow R^M$ is the objective function vector of the solution x.

Most often, the M objectives in Eq. (1) conflicts with each other. Hence, it is often impossible to achieve a solution that minimizes all the objectives simultaneously. In this case, the optimization goal becomes finding a set of tradeoff solutions with a good distribution. In detail, there are two main objectives when solving MOPs: minimizing the distance to the POF (i.e., convergence) and maximizing the spread along the POF (i.e., distribution).

After decades of research, MOEAs have demonstrated their potential for solving MOPs by providing a number of good findings. NSGA-II and SPEA2 are two most popular Pareto-based algorithms among early MOEAs. They use the Pareto-dominance relation as a primary selection criterion to drive a population of candidate solutions toward the POF and the distance-based density estimator as a second selection criterion to promote population diversity. The distance-based density estimator can in a sense remove overcrowded solutions from the population, but it cannot guarantee an even distribution of approximated points along the POF because of its insufficient estimation precision.

Decomposition-based algorithms are another class of MOEAs. This kind of algorithm employs effective decomposition approaches, such as the weighted sum approach or Tchebycheff approach, to convert an MOP into a number of single-objective problems and optimize them simultaneously. In such algorithms, a set of weight vectors helps to manage the distribution of approximated Pareto optimal points. A popular

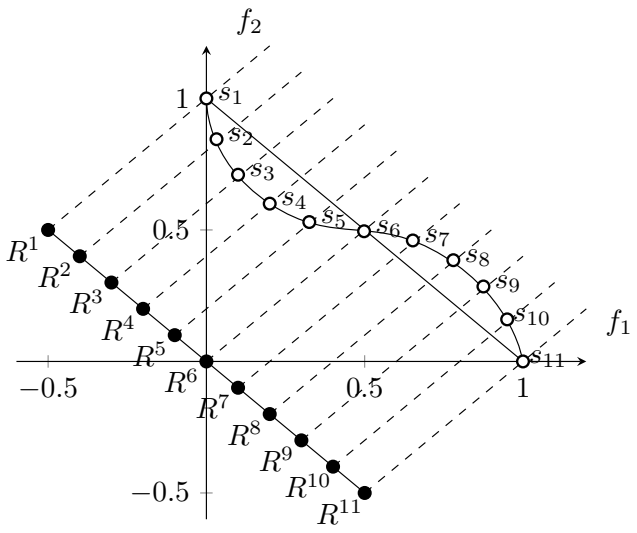

Fig. 1. Illustration of an even distribution of POF with the least reference distance (LRD).

hypothesis in decomposition-based algorithms is that an even distribution of weight vectors will result in well-distributed Pareto optimal points [21]. However, some recent studies showed that this assumption is fundamentally flawed [13], [7], [18]. The authors of [7] argue that an even or uniform distribution of weight vectors does not produce evenly distributed optimal solutions, thus effort to refine the distribution of weight vectors, such as the use of uniform design [20] and WS-transformation [18], does not fundamentally address this pressing issue. Even worse, decomposition-based algorithms struggle to evenly distribute Pareto optimal points along an excessively convex-shaped POF [13], [18], [11], where a small variation in one objective results in a large gap in another objective. These drawbacks mainly result from the use of decomposition approaches which nonlinearly map a set of uniformly-distributed weight vectors into the objective space [18], motivating many studies to explore more effective and powerful decomposition approaches [7], [11].

While both Pareto-based and decomposition-based methods struggle or even fail to achieve an even distribution of Pareto optimal points, the question then arises as to whether it is possible to fulfil this objective via other means. Inspired by decomposition-based methods, in this paper, we propose an LRD based MOEA (i.e., LRDA) for a good approximation of complex MOPs with irregular POF shapes.

Considering a normalized bi-objective POF (see Fig. 1), a line segment can be derived by joining two boundary points of the normalized POF, i.e., (0,1) and (1,0). Then it is easy to translate the line segment to a new position away from the normalized objective space (a unit square in the first quadrant), and we called the new line segment as reference line segment. Suppose that the reference line segment passes through the origin (actually, this is the minimum requirement for a reference line segment not intersecting the normalized objective space), and a set (R) of N reference points is evenly sampled from that line segment, then it is clear that any normal line through a sampled reference point intersects the normalized POF and all intersection points (denoted as $s_i, i = 1, \ldots, 11$ in

Fig. 1) are nearly as uniformly-distributed as reference points. Another fact is that an intersection point has only one nearest reference point with regard to the Euclidean distance. Thus, if a normalized point $\hat{F}(x)$ (normalized objective vector of x) wants to be an intersection point, it is required to move as close to its nearest reference point as possible. For this reason, we define a reference distance between x and a reference point R^i, $1 \leq i \leq N$, denoted as $g(x, R^i)$, as follows:

$$g(x, R^i) = \|\hat{F}(x) - R^i\| \tag{2}$$

where $\|\cdot\|$ denotes the Euclidean distance. Then, evolving x becomes identifying its nearest reference point j and at the same time minimizing the reference distance $g(x, R^j)$, which is actually the LRD that x can have. When x optimizes the LRD to the best of its ability, it is surely an intersection point and a Pareto optimal solution as well. This way, a number of intersection points can be found and their distribution is reasonably believed to be quite uniform.

It should be noted that, a similar idea was also introduced in the normal boundary intersection (NBI) [2], but LRDA mainly differs in its constraint-free property whereas NBI requires extra constraints. Another advantage of LRDA over NBI is that, unlike NBI, LRDA does not require the computation of the normal vector to the convex hull of individual minima (CHIM), thereby avoiding extra computational overheads.

III. Proposed LRD Algorithm (LRDA)

The general framework of LRDA is presented in Algorithm 1. LRDA starts with an initial population of N individuals randomly sampled from the search space and a set of evenly-distributed reference points generated from a shifted unit simplex. In each generational cycle, the ideal point is identified from the current population, which is used later for objective normalization. LRDA applies genetic operators to produce only one offspring solution at a time, and the nearest reference point to this solution is computed according to the LRD. Afterwards, this solution is further employed to update the individual associated with the nearest reference point. Generally speaking, LRDA is mainly composed of four key components, i.e., generation of reference set, genetic operator, objective normalization, and mating selection and solution replacement. In the following paragraphs, we give a detailed description of each component of the LRDA step by step.

A. Reference Set

A structured set of reference points can be generated spanning a hyperplane away from a normalized POF surface using Das and Dennis's systematic approach [2]. Specifically, the systematic approach first generates $N = \binom{p+M-1}{M-1}$ points on a unit simplex, with a uniform spacing $\delta_p = 1/p$, where p is the number of divisions along each objective coordinate. Then, the hyperplane is proportionally shifted away from its original position such that the sampled points on the translated hyperplane are not dominated by points in the non-negative orthant. Suppose that $\bar{R} = \{\bar{R}^1, \ldots, \bar{R}^N\}$, and δ_s are the original set sampled from the original unit simplex and the magnitude of shift in each objective coordinate, respectively, where all components of $\bar{R}^i$ satisfies $\sum_{j=1}^{M} \bar{R}^i_j = 1$, for any $1 \leq i \leq N$. The resulting set $R = \{R^1, \ldots, R^N\}$ after hyperplane shifting can be calculated by:

$$R^i_j = \bar{R}^i_j + \delta_s \tag{3}$$

To ensure that any point of the resulting reference set is not dominated by points in the non-negative orthant, the shifted hyperplane must not pass through any point in this area. This means, $\sum_{j=1}^{M} R^j_i \leq 0$ or simply $\delta_s \leq -1/M$ always holds. For simplicity, $\delta_s = -1/M$ is used in this paper, indicating that the shifted hyperplane passes exactly through the origin.

Algorithm 1 Framework of LRDA

1: **Input**:
   - a stopping criterion;
   - N: the number of reference points considered in LRDA;
2: **Output**: approximated Pareto-optimal set.
3: Create an initial population $P = \{x^1, \ldots, x^N\}$;
4: Sample N points from a unit simplex using Das and Dennis's systematic approach [2] and then shift them with a magnitude of $-\frac{1}{M}$ to generate the reference set $R = \{R^i, \ldots, R^N\}$;
5: **while** *stopping criterion not met* **do**
6: Compute intercepts $a = (a_1, \ldots, a_M)$ by Eq. (8);
7: **for** $i \leftarrow 1$ to N **do**
8: Set $r_1 = i$ and randomly select two indexes r_2 and r_3 from $\{1, \ldots, N\}$;
9: Apply the DE operator on individuals r_1, r_2 and r_3 by Eq. (4) to generate a solution $\bar{y}$, and perform the polynomial mutation operator on $\bar{y}$ by Eq. (5) to produce a new solution y;
10: Update the ideal point z^{min}: For each $j = 1, \ldots, M$, if $f_j(y) \leq z^{min}_j$, then set $z^{min}_j = f_j(y)$;
11: Compute the normalized objective vector of y according to Eq. (8);
12: Find the closest reference point to y: $\hat{j} = j$: argmin $g(y, R^j)$;
13: Update the $\hat{j}$th individual: If $g(y, R^{\hat{j}}) \leq g(x^{\hat{j}}, R^{\hat{j}})$, then set $x^{\hat{j}} = y$;
14: **end for**
15: **end while**
16: Output nondominated solution set.

B. Genetic Operations

In the reproduction process (lines 8-9 of Algorithm 1), any genetic operator can be used for producing a new offspring. For simplicity, we apply the differential evolution (DE) operator [17] and polynomial mutation operator [3] to produce offspring in LRDA, which is the case with MOEA/D-DE [15]. The DE operator generates a candidate solution $\hat{y}$ as follows:

$$\hat{y_k} = \begin{cases} x^{r_1}_k + F \times (x^{r_2}_k - x^{r_3}_k), & \text{with probability } CR, \\ x^{r_1}_k, & \text{with probability } 1\text{–}CR \end{cases} \tag{4}$$

where $\hat{y_k}$ is the kth component of $\hat{y}$, $x^{r_1}_k$, $x^{r_2}_k$, and $x^{r_3}_k$ are three distinct individuals randomly chosen from the population, and CR and F are two control parameters.

The polynomial mutation produces a solution $y = (y_1, \ldots, y_n)$ from $\hat{y}$ as follows:

$$y_k = \begin{cases} \hat{y_k} + \sigma_k \times (b_k - a_k), & \text{with probability } p_m, \\ \hat{y_k}, & \text{with probability } 1 - p_m \end{cases} \tag{5}$$

with

$$\sigma_k = \begin{cases} (2 \times rand)^{\frac{1}{\eta+1}}, & \text{if } rand < 0.5, \\ 1 - (2 - 2 \times rand)^{\frac{1}{\eta+1}}, & \text{otherwise} \end{cases} \quad (6)$$

where $rand$ is a uniform random number from [0, 1], the distribution index η and the mutation rate p_m are two control parameters, and b_k and a_k are the lower and upper bounds of the kth decision variable, respectively.

Note that, some MOEA/D variants define a neighbourhood size to select parents for mating, while the proposed algorithm considers the whole population as its mating range. Thus, LRDA does not require any extra parameter other than the usual genetic parameters, such as crossover and mutation related parameters, and is surely free from possible parameter sensitivity caused by additional parameters.

C. Objective Normalization

Objective normalization (line 11 of Algorithm 1) plays a crucial role in the uniformity of obtained nondominated points along the POF, that is, it can avoid possible biases (especially for disparately scaled objectives) when an algorithm pursues a set of evenly-distributed points on the POF, and its effectiveness has been verified in [21], [4]. Another important reason is that, in this paper, we consider the computation of reference distances (see Eq. (2)) in the normalized objective space. Like [4], the extreme point in each objective axis is identified by finding the solution that minimizing the following achievement scalarizing function:

$$ASF(x, w) = \max_{1 \le i \le M} \frac{f_i(x) - z^i_{min}}{w} \quad (7)$$

For identifying the extreme point in the ith objective axis, w should satisfy that $w_i = 1$ and the other components of w are set to 10^{-6}. This will generate M extreme points, which are later used to construct an M-dimensional linear hyperplane. Intercepts can be computed by finding the joining point of the constructed hyperplane and each coordinate axis. Based on these intercepts, the objective functions of an individual x can be normalized as follows:

$$\hat{f}_i(x) = \frac{f_i(x) - z^i_{min}}{a_i - z^i_{min}} \quad (8)$$

where $i \in \{1, \cdots, M\}$ and $\hat{f}_i(x)$ denotes the ith normalized objective of x. z^i_{min} and a_i are the ith component of the ideal point and the ith intercept, respectively.

D. Selection/Replacement

In the steady state form, if a child solution is better than one or more individuals in the population, it attempts to survive in the current population via a replacement. First, the child identifies its nearest reference point $R^{\hat{j}}$ in terms of the reference distance defined in Eq. (2). This aims to find the most possible POF segment or point toward which the child tries to converge. Then, it competes with the solution $x^{\hat{j}}$ associated with the jth reference point (the current best converging solution around that POF segment). If the child wins in terms of the reference distance, then it replaces $x^{\hat{j}}$. This selection and replacement procedure is presented in lines 12-13 of Algorithm 1.

IV. Experimental Study

A. Performance Metric

In our experiments, we mainly adopt the inverted generational distance (IGD) [26] as a performance indicator. IGD has been used since it can provide reliable information on both the diversity and convergence of obtained solutions. Let PF be a set of solutions uniformly sampled from the true POF, and PF^* be the approximated solutions in the objective space, the metric measures the gap between PF^* and PF, which is calculated as follows:

$$IGD(PF^*, PF) = \frac{\sum_{p \in PF} d(p, PF^*)}{|PF|} \quad (9)$$

where $d(p, PF^*)$ is the distance between the member p of PF and the nearest member of PF^*.

B. Experimental Setting

As our aim is to achieve an even distribution of obtained solutions for complex problems, the test problems chosen for comparison should contain complex characteristics, especially excessively convex-shaped fronts. Besides borrowing three test problems from [11], we also constructed two new instances with mixed (concave and convex) POF shapes. These test problems are detailed in Table I, and their POFs are known *a prior* so that the IGD metric can be easily computed.

Our proposed LRDA algorithm is compared with three popular algorithms, i.e., NSGA-II [5], MOEA/D-TCH [21], and MOEA/D-PBI [21]. MOEA/D-TCH and MOEA/D-PBI are two MOEA/D variants using the Tchebycheff approach and the penalty boundary intersection (PBI) approach, respectively. To make a fair comparison, all the tested algorithms used the DE operator for recombination. Two control parameters in DE were $CR = 1.0$ and $F = 0.5$. The mutation rate was $p_m = 1/n$ and its distribution $\eta_m = 20$. The neighbourhood size T in two MOEA/D variants was set to 20. According to [15], with a probability of 1-δ to choose the whole population as neighbourhood, MOEA/D variants appear to improve the population diversity. Additionally, the maximal number n_r of solutions replaced by a child solution was suggested to be much smaller than T. For these reasons, here we set $\delta = 0.9$ and $n_r = 2$, as recommended in [15].

The population size was set to 100 and the maximum number of generations 300. For each test instance, each algorithm was independently executed 31 runs.

C. Experimental Results

Table II presents the IGD values obtained by four algorithms for five test problems, where the best values are highlighted in bold face. To have a clear and graphical understanding of the performance of these compared algorithms, we also plot their approximated POFs corresponding to the median IGD metric in Fig. 2.

F1 has a continuous POF, but its POF geometry is excessively convex-shaped, which poses a considerable challenge to algorithms' distribution on boundary regions. From the perspective of the IGD metric, it is clear that LRDA performs the best, followed by NSGA-II, and neither of the MOEA/D variants provides satisfactory IGD values in this case. This

TABLE I. TEST INSTANCES

Instance	Description	Domain	Number of Variables	Notes
F1	$f_1(x) = (1 + g(x))x_1$ $f_2(x) = (1 + g(x))(1 - \sqrt{x_1})^5$ $g(x) = 2\sin(0.5\pi x_1)(n - 1 + \sum_{i=2}^{n}(y_i^2 - \cos(2\pi y_i)))$ where $y_{i=2:n} = x_i - \sin(0.5\pi x_i)$ POF: $f_2 = (1 - \sqrt{f_1})^5$ POS: $x_i = \sin(0.5\pi x_i), i = 2, \ldots, n$	$[0, 1]^n$	10	Uni-modal Convex Separable
F2	$f_1(x) = (1 + g(x))(1 - x_1)$ $f_2(x) = \frac{1}{2}(1 + g(x))(x_1 + \sqrt{x_1}\cos^2(4\pi x_1))$ $g(x) = 2\sin(0.5\pi x_1)(n - 1 + \sum_{i=2}^{n}(y_i^2 - \cos(2\pi y_i)))$ where $y_{i=2:n} = x_i - \sin(0.5\pi x_i)$ POF: $f_2 = \frac{1}{2}(1 - f_1 + \sqrt{1 - f_1}\cos^2(4\pi(1 - f_1)))$ POS: $x_i = \sin(0.5\pi x_i), i = 2, \ldots, n$	$[0, 1]^n$	10	Multi-modal Disconnected Separable
F3	$f_1(x) = (1 + g(x))x_1$ $f_2(x) = \frac{1}{2}(1 + g(x))(1 - x_1^{0.1} + (1 - \sqrt{x_1})^2\cos^2(3\pi x_1))$ $g(x) = 2\sin(0.5\pi x_1)(n - 1 + \sum_{i=2}^{n}(y_i^2 - \cos(2\pi y_i)))$ where $y_{i=2:n} = x_i - \sin(0.5\pi x_i)$ POF: $f_2 = \frac{1}{2}(1 - f_1^{0.1} + (1 - \sqrt{f_1})^2\cos^2(3\pi f_1))$ POS: $x_i = \sin(0.5\pi x_i), \forall x_i \in \mathbf{x_{II}}$	$[0, 1]^n$	10	Multi-modal Disconnected Separable
F4	$f_1(x) = (1 + g(x))(x_1 + 0.05\sin(6\pi x_1))^2$ $f_2(x) = (1 + g(x))(1 - x_1 + 0.05\sin(6\pi x_1))^2$ $g(x) = 2\sin(0.5\pi x_1)(n - 1 + \sum_{i=2}^{n}(y_i^2 - \cos(2\pi y_i)))$ where $y_{i=2:n} = x_i - \sin(0.5\pi x_i)$ POF: $f_1^{0.5} + f_2^{0.5} = 1 + 0.1\sin(3\pi(f_1^{0.5} - f_2^{0.5} + 1))$ POS: $x_i = \sin(0.5\pi x_i), \forall x_i \in \mathbf{x_{II}}$	$[0, 1]^n$	10	Multi-modal mixed Separable
F5	$f_1(x) = (1 + g(x))(x_1 + 0.05\sin(6\pi x_1))^{20}$ $f_2(x) = (1 + g(x))(1 - x_1 + 0.05\sin(6\pi x_1))^{0.2}$ $g(x) = 2\sin(0.5\pi x_1)(n - 1 + \sum_{i=2}^{n}(y_i^2 - \cos(2\pi y_i)))$ where $y_{i=2:n} = x_i - \sin(0.5\pi x_i)$ POF: $f_1^{0.05} + f_2^5 = 1 + 0.1\sin(3\pi(f_1^{0.05} - f_2^5 + 1))$ POS: $x_i = \sin(0.5\pi x_i), \forall x_i \in \mathbf{x_{II}}$	$[0, 1]^n$	10	Multi-modal mixed Non-separable

TABLE II. BEST, MEDIAN AND WORST IGD VALUES OF FOUR ALGORITHMS FOR FIVE TEST PROBLEMS

Problem	NSGA-II	MOEA/D-TCH	MOEA/D-PBI	LRDA
F1	5.4270E-03	4.1330E-02	2.0611E-01	**4.5650E-03**
	5.8190E-03	4.2585E-02	2.0749E-01	**4.5760E-03**
	2.0843E-01	5.5038E-02	2.0821E-01	**4.6730E-03**
F2	**3.3090E-03**	4.5140E-03	1.8888E-02	3.9830E-03
	3.8005E-03	4.9310E-03	2.0027E-02	4.0855E-03
	1.5573E-01	1.5770E-01	2.1835E-02	**4.6730E-03**
F3	**3.9410E-03**	1.8197E-02	1.9520E-01	4.4870E-03
	4.3950E-03	1.9399E-02	3.9850E-01	4.5605E-03
	4.7320E-03	2.0621E-02	3.9879E-01	9.3520E-03
F4	5.2040E-03	1.2622E-02	8.3276E-02	**4.4570E-03**
	5.5375E-03	1.2689E-02	8.3769E-02	**4.4865E-03**
	6.2610E-03	1.2737E-02	8.4260E-02	**4.6140E-03**
F5	4.9870E-03	1.3534E-02	4.6824E-02	**4.2390E-03**
	5.4870E-03	1.4371E-02	6.0063E-02	**4.3200E-03**
	5.9400E-03	1.4807E-02	7.4753E-01	**5.0650E-03**

conclusion is further confirmed by the approximated POFs of four algorithms, where both NSGA-II and LRDA completely cover the POF, and LRDA distributes points more uniformly than NSGA-II along the POF. MOEA/D-TCH and MOEA/D-PBI, however, favour the intermediate region of the POF and thus miss boundary parts of the POF.

F2 is a disconnected problem, whose POF has several discontinuous regions. It challenges algorithms in finding all POF regions and maintaining diversity in each found POF region. For this problem, NSGA-II performs slightly better than LRDA in terms of the IGD metric, but it performs poorly in some runs, as indicated by a large IGD value. The IGD values given by two MOEA/D variants again are not encouraging. The obtained POF plots show that four algorithms are all able to find the whole POF regions, and LRDA appears to provide a more uniform distribution of Pareto optimal points than the others.

Besides disconnectivity, F3 also has an excessively convex POF shape. Thus, it is much harder than F2. For F3, the IGD values obtained by NSGA-II are slightly better than those by LRDA, and the IGD values also suggest both MOEA/D variants have difficulties in handling this kind of problem. As can be seen from their approximated POFs, both NSGA-II and LRDA are able to offer a good approximation for this case, and LRDA gives the best distribution. In contrast, the performance of both MOEA/D variants is not impressive, although MOEA/D-TCH finds a large majority of Pareto optimal points along the POF and MOEA/D-PBI provides many solutions in the intermediate region. Both of them fail to obtain uniformly distributed solutions on each disconnected POF region.

Problems F4 and F5 have mixed and excessively-convex POF shapes. For these two problems, LRDA obtains the best IGD values. The IGD values obtained by NSGA-II are also competitive, although they are slightly worse than those obtained by LRDA. MOEA/D-PBI performs significantly worse than MOEA/D-TCH in terms of the IGD metric. Considering the approximated POFs, it is evident that only LRDA is able to produce an extremely even distribution of Pareto optimal points for both problems. MOEA/D-PBI misses points on boundary regions whereas MOEA/D-TCH cannot distribute

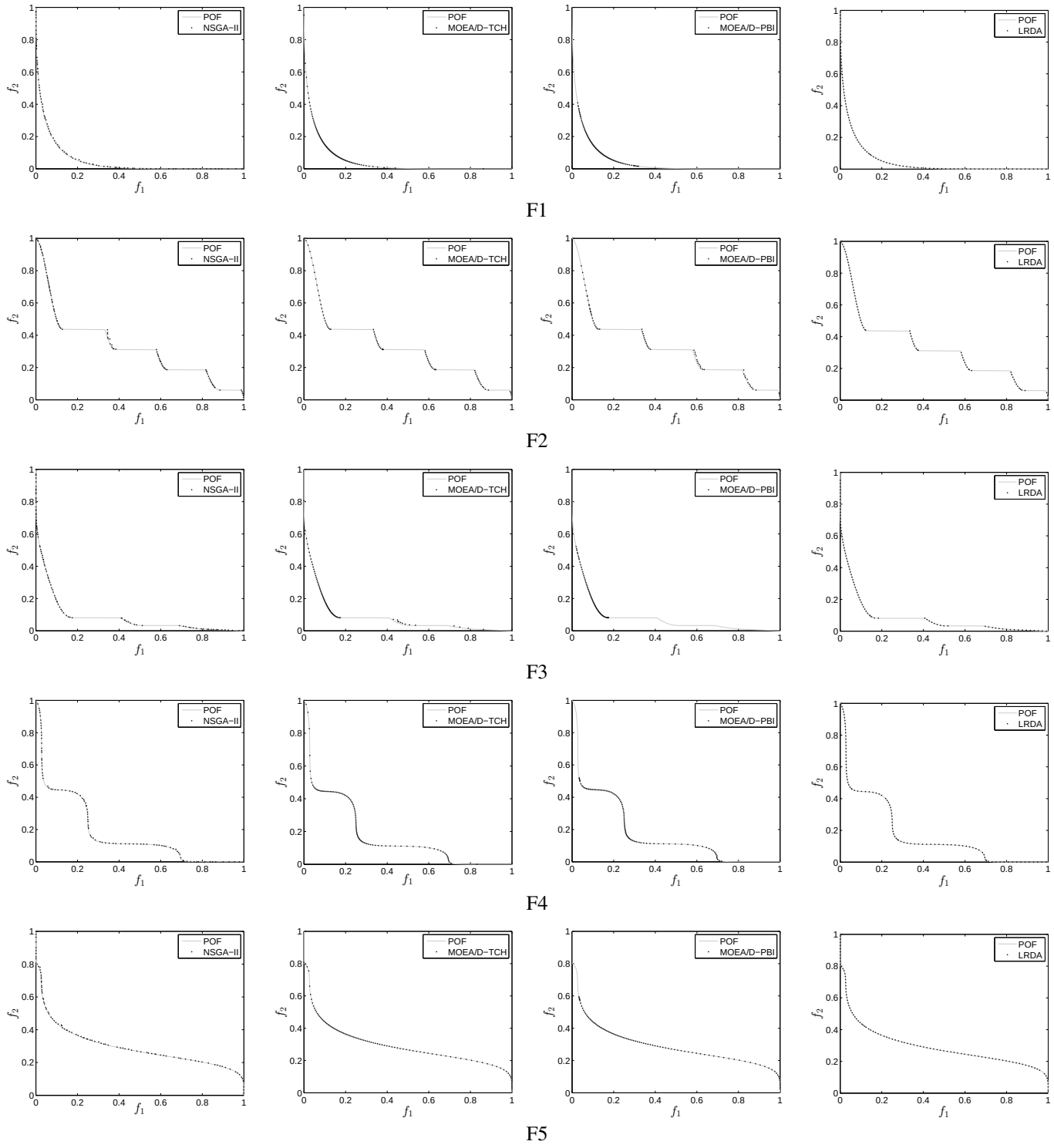

Fig. 2. Approximated POFs obtained by for algorithms for five test problems.

approximated points evenly along the POF.

It is reasonable that LRDA does not outperform NSGA-II on two disconnected problems, i.e., F2 and F3, in terms of the IGD metric. If the true POF to be approximated is discontinuous, LRDA will generate some dominated solutions between two adjoining disconnected POF regions. This means a population of N individuals will produce less than N Pareto optimal points, which are then used for IGD computation. In contrast, under the same condition, NSGA-II probably produces exact N Pareto optimal points. The more Pareto optimal points an algorithm obtains, the better the IGD metric will probably be. Thus, NSGA-II gives slightly better IGD values than LRDA. On the other hand, both MOEA/D variants perform poorly on all the test problems. One reason for their pool performance is that, due to irregular POF shapes, the

test problems used in this paper are much more complex than some commonly-used test suites like ZDT [23], for which the study of [21] has reported that the MOEA/D variants outperform NSGA-II. In this work, however, the MOEA/D variants perform evidently worse than NSGA-II. The non-uniformity of approximated points obtained by two MOEA/D variants also suggest that the Tchebycheff approach or the PBI approach might not be suitable for solving MOPs with complex POFs considered in this paper and the utility of decomposition approaches for this kind of problem should be further explored.

V. Conclusions and Future Work

A novel LRDA algorithm for multiobjective optimization has been presented in this paper. The LRDA works with a set of predefined and uniformly-distributed reference points and steadily update solutions associated with reference points that are closest to a new solution during the evolution. Through minimizing the LRD metric, a population of solutions are hopeful to be as close to the reference set as possible, thereby providing an even distribution of Pareto optimal points. The proposed algorithm has been tested and compared with several popular algorithms on a number of old and newly-constructed MOPs with complex Pareto fronts, and the experimental results have shown that LRDA is able to approximate the complex POFs with a uniform distribution, and in some problems, LRDA even outperforms NSGA-II and two MOEA/D variants. This demonstrates that LRDA is very promising for multiobjective optimization.

Although it has offered encouraging results on the test problems considered in this paper, the proposed algorithm should be further investigated in a wider range of test problems, including MOPs with complicated POSs or high-dimensional problems. As discussed in the experimental study, LRDA is likely to generate some dominated solutions for discontinuous problems. Thus, it will be interesting to introduce some techniques to reduce this drawback, which is left as our future work.

Acknowledgment

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) of UK under Grant EP/K001310/1.

References

[1] D. W. Corne and J. D. Knowles, "The pareto-envolope based selection algorithm for multiobjective optimization," in *Proc. 6th Int. Conf. Parallel Problem Solving from Nature (PPSN VI)*, vol. 1917, 2000, pp. 839–848.

[2] I. Das and J. Dennis, "Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems," *SIAM J. Optimiz.*, vol. 8, no. 3, pp. 631–657, 1998.

[3] K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms*. Chichester: John Wiley and Sons, 2001.

[4] K. Deb and H. Jain, "An evolutionary many-objective optimization algorithm using reference-point based non-dominated sorting approach, Part I: Solving problems with box constraints," *IEEE Trans. Evol. Comput.*, vol. 18, no. 4, pp. 577–601, 2014.

[5] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, 2002.

[6] C. M. Fonseca and P. J. Fleming, " Genetic algorithms for multiobjective optimization: Formulation, discussion and generalization," in *Proc. 5th Int. Conf. Genetic Algorithms*, 1993, pp. 456–423.

[7] I. Giagkiozis, R. C. Purshouse, and P. J. Fleming, "Generalized decomposition and cross entropy methods for many-objective optimization," *Inform. Sci.*, vol. 282, pp. 363–387, Oct. 2014.

[8] J. Horn, N. Nafpliotis, and D. E. Goldberg, "A nched pareto genetic algorithm for multiobjective optimization," in *Proc. 1st IEEE Conf. Evol. Comput.*, vol. 1, 1994, pp. 82–87.

[9] H. Ishibuchi, Y. Sakane, N. Tsukamoto, and Y. Nojima, "Adaptation of scalarizing functions in MOEA/D: An adaptive scalarizing function-based multiobjective evolutionary algorithm," in *Proc. 5th Int. Conf. Evol. Multi-Criterion Optim.*, 2009, pp. 438–452.

[10] H. Ishibuchi, Y. Sakane, N. Tsukamoto, and Y. Nojima, "Simultaneous use of different scalarizing functions in MOEA/D," in *Proc. 2010 Genetic Evol. Comput. Conf.*, 2010, pp. 519–526.

[11] S. Jiang and S. Yang, "An improved multi-objective optimization evolutionary algorithm based on decomposition for complex Pareto fronts," *IEEE Trans. Cybern.*, 2015, in Press. Available: DOI: 10.1109/TCYB.2015.2403131.

[12] J. D. Knowles and D. W. Corne, "Approximating the nondominated front using the pareto archived evolution strategy," *Evol. Comput.*, vol. 8, pp. 149–172, 1999.

[13] H. Liu, F. Gu, and Y. Cheung, "T-MOEA/D: MOEA/D with objective transform in multiobjective problems," in *Proc. Int. Conf. Inform. Sci. and Manage. Eng.*, 2010, pp. 282–285.

[14] H. Liu, F. Gu, and Q. Zhang, "Decomposition of a multiobjective optimization problem into a number of simple multiobjective subproblems," *IEEE Trans. Evol. Comput.*, vol. 18, no. 3, pp. 450–455, Jun. 2014.

[15] H. Li and Q. Zhang, "Multiobjective optimization problems with complicated Pareto sets, MOEA/D and NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 13, no. 2, pp. 284–302, Apr. 2009.

[16] K. Miettnen, *Noliner Multiobjective Optimization*. New York, NY, USA: Springer, vol. 12, 1999.

[17] K. Price, R. M. Storn, and J. A. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization* (Natural Computing). Secaucus, NJ, USA: Springer, 2005.

[18] T. Qi, X. Ma, F. Liu, L. Jiao, J. Sun, and J. Wu, "MOEA/D with adaptive weight adjustment," *Evol. Comput.*, vol. 22, no. 2, pp. 231–264, 2014.

[19] N. Srinivas and K. Deb, "Multiobjective optimization using nondominated sorting in genetic algorithms, *Evol. Comput.*, vol. 2, no. 3, pp. 221–248, 1994.

[20] Y. Tan, Y. Jiao, H. Li, and X. Wang, "MOEA/D+uniform design: A new version of MOEA/D for optimization problems with many objectives," *Comput. Oper. Res.*, vol. 40, no. 6, pp. 1648–1660, 2013.

[21] Q. Zhang and H. Li, "MOEA/D: A multiobjective evolutionary algorithm based on decomposition," *IEEE Trans. Evol. Comput.*, vol. 11, no. 6, pp. 712–731, 2007.

[22] Q. Zhang, W. Liu, and H. Li, "The performance of a new version of MOEA/D on CEC09 unconstrained MOP test instances," School Comput. Sci. Electr. Eng., Univ. of Essex, U.K., Tech. Rep. CES-491, 2009.

[23] E. Zitzler, K. Deb, and L. Thiele, "Comparison of multiobjective evolutionary algorithms: empirical results," *Evol. Comput.*, vol. 8, no. 2, pp. 173–195, 2000.

[24] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: Improving the strength Pareto evolutionary algorithm for multiobjective optimization," in *Proc. Evol. Methods Des. Optimiz. Contr. Applicat. Indust. Problems*, EUROGEN, vol. 3242, no. 103, 2002, pp. 95-100.

[25] E. Zitzler and Thiele, "Multiobjective evolutionary algorithms: A comparative case study and the strength pareto approach," *IEEE Trans. Evol. Comput.*, vol. 3, no. 4, pp. 257–271, 1999.

[26] E. Zitzler, L. Thiele, M. Laumanns, C.M. Fonseca, and V. G. da Fonseca, "Performance assessment of multiobjective optimizers: An analysis and review," *IEEE Trans. Evol. Comput.*, vol. 7, no. 2, pp. 117–132, 2003.